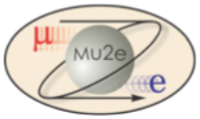

Mu2e-doc-1794-v2

Mu2e Fermigrid Incident of July 15, 2011

Rob Kutschke, Fermilab

July 25, 2011

Our Apologies

- From Mu2e to those whose jobs we affected and to those who had to break away from their regular work to diagnose this problem.

What Happened

- On July 15, 2011 a Mu2e user was performing a run of MC events.
- One cluster with 200 processes. Each process:
  - 50,000 events
  - 3 hours
  - 10 GB
- All well within guidelines.
- Had been tested with a lower statistics run.
- He fat-fingered an extra 0 into the number of events per job ….
- This caused overflows of the worker node disks which killed his jobs and other jobs working on the same nodes.
- When jobs exited the disk space was not cleaned up so subsequent jobs starting on those nodes also failed.

Stopping This in the Future

- Inform users that we can write 50E3 or 5E4 for the number of events instead of 50000.
  - Not really a solution but …
- We have asked the art developers to develop tools to observe total file size of output files and shutdown gracefully when a threshold is crossed.
  - We will have a design discussion but this is not yet scheduled.
- We have some options to reduce output file size
  - By default we keep the full G4 history.
  - We envisage weeding this of branches that do not lead to interesting leaves.
  - I don't know if this is relevant for this particular problem.
- What do others do?

We do Auto-delete files

```
ORIGDIR=`pwd`
TMP=`mktemp -d ${OSG_WN_TMP:-/var/tmp}/working_dir.XXXXXXXXXX`
TMP=${TMP:-${OSG_WN_TMP:-/var/tmp}/working_dir.$$}

{ [[ -n "$TMP" ]] && mkdir -p "$TMP"; } || \
  { echo "ERROR: unable to create temporary directory!" 1>&2; exit 1; }
trap "[[ -n \"$TMP\" ]] && { cd ; rm -rf \"$TMP\"; }" 0
cd $TMP

if (( $( ls -1 $ORIGDIR | wc -l) > 0 )); then
 mv $ORIGDIR/* .
fi

# User code starts here.
```

For jobs that died with full disks, the trap failed. Not understood.